

On uniformly generating Latin squares

M. Aryapoor* and E. S. Mahmoodian†

Abstract

By simulating an ergodic Markov chain whose stationary distribution is uniform over the space of $n \times n$ Latin squares, Mark T. Jacobson and Peter Matthews [4], have discussed elegant methods by which they generate Latin squares with a uniform distribution (approximately). The central issue is the construction of “moves” that connect the squares. Most of their lengthy paper is to prove that the associated graph is indeed connected. We give a short proof of this fact by using the concepts of Latin bitrades.

1 Introduction and preliminaries

A **Latin square** L of order n is an $n \times n$ array with entries chosen from an n -set N , e.g. $\{1, \dots, n\}$, in such a way that each element of N occurs precisely once in each row and column of the array. A **partial Latin square** P of order n is an $n \times n$ array with entries chosen from an n -set N , in such a way that each element of N occurs at most once in each row and at most once in each column of the array. Hence there are cells in the array that may be empty, but the positions that are filled have been so as to conform with the Latin property of array. For ease of exposition, a partial Latin square T may be represented as a set of ordered triples: $\{(i, j; T_{ij}) \mid \text{where element } T_{ij} \text{ occurs in (nonempty) cell } (i, j) \text{ of the array}\}$.

Let T be a partial Latin square and L a Latin square such that $T \subseteq L$. Then T is called a **Latin trade**, if there exists a partial Latin square T^* such that $T^* \cap T = \emptyset$ and $(L \setminus T) \cup T^*$ is a Latin square. We call T^* a **disjoint mate** of T and the pair $\mathcal{T} = (T, T^*)$ is called a **Latin bitrade**. The **volume** of a Latin bitrade is the number of its nonempty cells. A Latin bitrade of volume 4 which is unique (up to isomorphism), is said to be an **intercalate**. A bitrade $\mathcal{T} = (T, T^*)$ may be viewed as a set of positive triples T and negative triples T^* .

*Institute for Studies in Theoretical Physics and Mathematics (IPM), Niavaran Square, Tehran, Iran (masood.aryapoor@ipm.ir).

†Department of Mathematical Sciences, Sharif University of Technology, P. O. Box 11155-9415, Tehran, Iran (emahmood@sharif.edu).

Example 1 The bitrade $\mathcal{I} = (I, I^*)$, where

$$I = \{(i, j; a), (i, j'; b), (i', j; b), (i', j'; a)\},$$

$$I^* = \{(i, j; b), (i, j'; a), (i', j; a), (i', j'; b)\},$$

is an intercalate. Note that we must have $i \neq i'$, $j \neq j'$ and $a \neq b$. Usually such an intercalate is shown as

		j	j'	
		\cdot	\cdot	\cdot
\mathcal{I}	i	a_b	b_a	\cdot
		\cdot	\cdot	\cdot
	i'	b_a	a_b	\cdot
		\cdot	\cdot	\cdot

where the elements of I^* are written as subscripts in the same array as I .

For a recent survey on Latin bitrades see [2] and also [5].

In [4] the approach for generating Latin squares is based on the fact that an $n \times n$ Latin square is equivalent to an $n \times n \times n$ contingency (proper) table in which each line sum equals 1. They relax the nonnegativity condition on the table's cells, allowing "improper" tables that have a single -1 -cell. A simple set of moves connects this expanded space of tables [the diameter of the associated graph is bounded by $2(n-1)^3$] and suggests a Markov chain whose subchain of proper tables has the desired uniform stationary distribution. By grouping these moves appropriately, they derive a class of moves that stay within the space of proper Latin squares.

An improper Latin square is an $n \times n$ array such that each cell has a single symbol, except for *one* improper cell (in the improper row and column) which has three (the improper symbol appears there with a -1 coefficient). Each symbol appears exactly once in each row and in each column, except in the improper row (and also in the improper column) where one of the symbols appears twice as "positive" and once as "negative". An improper Latin square may be viewed as a set of $n^2 + 1$ positive triples and one negative triple.

Example 2 The following array is an improper Latin square of order 4.

	c	b	d	a
	b	d	a	c
L	d	$a + c - b$	b	b
	a	b	c	d

Using the notation of Latin bitrades, we may show this improper Latin square by

$$L \begin{array}{cccc} c & b & d & a \\ b & d & a & c \\ d & a + c_b & b & b \\ a & b & c & d \end{array}$$

The notion of ± 1 -move is introduced in [4]. Using the notation of Latin bitrades, a ± 1 -move means adding some appropriate intercalate to a given proper or improper Latin square such that the result is a proper or improper Latin square. If the added intercalate is the intercalate \mathcal{I} in Example 1, the corresponding ± 1 -move is called a $((i, j; a), (i', j'; b))$ -move.

Example 3 By applying the $((1, 2; a), (3, 4; b))$ -move to the improper Latin square L in Example 2, we obtain the following Latin square.

$$L' \begin{array}{cccc} c & a & d & b \\ b & d & a & c \\ d & c & b & a \\ a & b & c & d \end{array}$$

Let $G = (V, E)$ be a graph whose vertices are associated to S , the set of all proper and improper Latin squares of order n , and two vertices L and L' are adjacent if there is a ± 1 -move transferring L to L' . In the next section we state the results which prove that G is connected. This approach is developed from a linear algebraic approach to the concept of Latin bitrades, which is detailed in the references [6], [7] and [3].

2 Connectivity of graph G

In this section we prove that the graph G (defined in the last section) is connected. First we need a few lemmas. The first lemma states that an improper Latin square can be transferred into a proper Latin square using ± 1 -moves with changes only in two rows.

Lemma 1 Suppose that we have the following improper Latin square

$$A \begin{array}{c|cccc} & & j & & \\ \hline & \cdot & \cdot & \cdot & \\ i_1 & \cdot & a + b_s & \cdot & \\ & \cdot & \cdot & \cdot & \\ i_2 & \cdot & s & \cdot & \\ & \cdot & \cdot & \cdot & \end{array}$$

Then there is a sequence of (at most $\frac{n-1}{2}$) ± 1 -moves involving only rows i_1 and i_2 which transfers A to a proper Latin square.

Proof. It is easy to see that we can find the following cyclic pattern lying in rows i_1 and i_2 of A (possibly after permuting some columns of A)

		j	j_1	j_2	j_3	\cdots	j_{r-1}	j_r	
A	i_1	\cdot	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot	\cdot
		\cdot	$a + b_s$	t	u	v	\cdots	z	s
	i_2	\cdot	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot	\cdot
		\cdot	s	b	t	u	\cdots	y	z

where $t, u, \dots, z \notin \{s, a, b\}$ or $r = 1$ (i.e. $t = s$). Note that there is a similar pattern corresponding to a which has no intersection with the above pattern except in the j th column. Therefore one of these patterns is at most of length $\frac{n-1}{2}$, and we may assume that $r \leq \frac{n-1}{2}$. We proceed by induction on r . If $r = 1$, then the $((i_1, j; s), (i_2, j_1; b))$ -move produces a proper Latin square. Let $r > 1$. Then the $((i_1, j; s), (i_2, j_r; b))$ -move decreases r . ■

In the next lemma we show that one can swap a cycle lying in two rows using ± 1 -moves. In [9] it is called a cycle switch.

Lemma 2 Suppose we have the following cyclic pattern in a (proper) Latin square

		j_1	j_2	j_3	\cdots	j_{r-1}	j_r	
A	i_1	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot	\cdot
		\cdot	s	t	u	y	z	\cdot
	i_2	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot	\cdot
		\cdot	t	u	v	\cdots	z	s

Then there is a sequence of (exactly $r - 1$) ± 1 -moves acting only on the entries shown above which transfers A to

		j_1	j_2	j_3	\cdots	j_{r-1}	j_r	
	i_1	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot	\cdot
		\cdot	t	u	v	\cdots	z	s
	i_2	\cdot	\cdot	\cdot	\cdots	\cdot	\cdot	\cdot
		\cdot	s	t	u	\cdots	y	z

Proof. If $r = 2$ (i.e. $u = s$), then the $((i_1, j_1; t), (i_2, j_2; s))$ -move does the job. So let $r \geq 3$. Then the $((i_1, j_1; t), (i_2, j_2; s))$ -move transfers A to

		j_1	j_2	j_3	\cdots	j_{r-1}	j_r	
	i_1	\cdot	t	s	u	\cdots	y	z
	i_2	\cdot	s	$u + t_s$	v	\cdots	t	s

Now by applying the method in the proof of Lemma 1, this improper Latin square can be transferred to the desired Latin square. ■

The last lemma is a crucial lemma. It tells us that we can switch two entries in a row of an improper Latin square using a sequence of controlled ± 1 -moves.

Lemma 3 *Suppose for given s and $t \in \{1, 2, \dots, n\}$ we have the following improper Latin square:*

		j_1	j_2	
	i_1	\cdot	s	t
	i_2	\cdot	$a + b_s$	\cdot
	i_3	\cdot	\cdot	s

where i_2 may be equal to i_3 . Then there is a sequence of (at most $2(n-1)$) ± 1 -moves transferring this square to an improper (or proper) Latin square A' of the following form:

		j_1	j_2	
	i_1	\cdot	t	s
	i'	\cdot	$e + f_t$	\cdot

where $i' = i_2$ or i_3 , and the only possibly different entries of A and A' are entries in: (i_1, j_1) , (i_1, j_2) and those in rows i_2 and i_3 .

Proof. We distinguish two cases.

Case 1: $i_2 = i_3$, i.e. in column j_2 the symbol s appears in the improper row. So A has the following form:

		j_1	j_2
A	i_1	\cdot	\cdot
	i_2	$a + b_s$	s
		\cdot	\cdot
		\cdot	\cdot

Then the $((i_1, j_1; t), (i_2, j_2; s))$ -move transfers A to:

		j_1	j_2
A'	i_1	\cdot	\cdot
	i_2	$a + b_t$	t
		\cdot	\cdot
		\cdot	\cdot

and we are done.

Case 2: $i_2 \neq i_3$.

It is easy to see that we can find the following cyclic pattern lying in rows i_2 and i_3 of A (possibly after permuting some columns of A)

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2
A	i_1	\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot
	i_2	s	c	\cdot	\cdot	\dots	\cdot	\cdot	t
	i_3	$a + b_s$	s	u	v	\dots	x	y	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot
		d	u	v	w	\dots	y	z	s
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot

where $\{u, v, w, \dots, x, y\} \cap \{a, b\} = \emptyset$, but $z \in \{a, b\}$. Without loss of generality we assume that $z = b$. Therefore A has the following cyclic pattern

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2
A	i_1	\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot
	i_2	s	c	\cdot	\cdot	\dots	\cdot	\cdot	t
	i_3	$a + b_s$	s	u	v	\dots	x	y	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot
		d	u	v	w	\dots	y	b	s
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot

Then the $((i_2, j_1; s), (i_1, c_1; b))$ -move transfers A to:

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2	
A_1	i_1	\cdot	b	$c + sb$	\cdot	\dots	\cdot	\cdot	t	\cdot
	i_2	\cdot	a	b	u	v	\dots	x	y	\cdot
	i_3	\cdot	d	u	v	w	\dots	y	b	s
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot

If $u = b$ (i.e. $r = 1$) then the $((i_3, j_1; b), (i_1, c_1; s))$ -move transfers A_1 to:

	j_1	c_1	j_2	
i_1	\cdot	s	c	t
i_2	\cdot	a	b	\cdot
i_3	\cdot	$d + bs$	s	s

which reduces the problem to Case 1. So we assume that $u \neq b$. Now the symbol b appears once more as a positive entry in column c_1 and another row, say i_4 :

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2	
A_1	i_1	\cdot	b	$c + sb$	\cdot	\dots	\cdot	\cdot	t	\cdot
	i_2	\cdot	a	b	u	v	\dots	x	y	\cdot
	i_3	\cdot	d	u	v	w	\dots	y	b	s
	i_4	\cdot	\cdot	b	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot

where $i_4 \notin \{i_1, i_2, i_3\}$. By Lemma 1, with a sequence of ± 1 -moves only on rows i_1 and i_4 , we can obtain a proper Latin square A_2 . Using Lemma 2, a sequence of ± 1 -moves transfers A_2 to

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2	
A_3	i_2	\cdot	a	u	v	w	\dots	y	b	\cdot
	i_3	\cdot	d	b	u	v	\dots	x	y	s
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot

Now we can undo the sequence of ± 1 -moves on rows i_1 and i_4 in A_3 , to obtain the corresponding rows in A_1 . The resulting Latin square has the following pattern

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2	
A_4	i_1	\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
		\cdot	b	$c + sb$	\cdot	\dots	\cdot	\cdot	t	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
	i_2	\cdot	a	u	v	\dots	y	b	\cdot	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
	i_3	\cdot	d	b	u	\dots	x	y	s	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot

With the $((i_3, j_1; b), (i_1, c_1, s))$ -move, A_4 can be transferred to

		j_1	c_1	c_2	c_3	\dots	c_{r-1}	c_r	j_2	
A_5	i_1	\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
		\cdot	s	c	\cdot	\dots	\cdot	\cdot	t	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
	i_2	\cdot	a	u	v	\dots	y	b	\cdot	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot
	i_3	\cdot	$d + bs$	s	u	\dots	x	y	s	\cdot
		\cdot	\cdot	\cdot	\cdot	\dots	\cdot	\cdot	\cdot	\cdot

and finally Case 1 finishes the proof. Note that in row i_1 , except the positions of s and t , other positions are unchanged. \blacksquare

Now we can prove that the graph G is connected.

Theorem 1 *Let S be the set of all proper or improper Latin squares of order n . Given two Latin squares of order n , there exists a sequence of ± 1 -moves that transfers one square into the other without leaving S . An upper bound on the length of the shortest such sequence is $2(n-1)^3$.*

Proof. Suppose that A and B are two proper or improper Latin squares. Without loss of generality we can assume that A and B are proper (see Lemma 1). To prove the theorem, we proceed by induction on the number of identical rows of A and B . Suppose that the first $k-1$ rows of A and B are equal. We show that we can apply a sequence of ± 1 -moves to A to obtain a Latin square with the first k rows identical to the first k rows of B . If the k th rows of A and B are equal then we are done. So suppose that they are not equal. In this case we can find the following patterns in A and B ($s \neq a$)

$$A = \begin{array}{c|ccccc} & & j_1 & & j_2 & \\ \hline & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & \cdot & s & \cdot & t & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_1 & \cdot & u & \cdot & s & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \quad B = \begin{array}{c|ccccc} & & j_1 & & j_2 & \\ \hline & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & \cdot & a & \cdot & s & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_1 & \cdot & b & \cdot & c & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

Since A and B are proper Latin squares and have the same first $k - 1$ rows, we must have $i_1 > k$. Now the $((k, j_2; s), (i_1, j_1; t))$ -move transfers A to

$$\begin{array}{c|ccccc} & & j_1 & & j_2 & \\ \hline & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & \cdot & t & \cdot & s & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_1 & \cdot & u + s_t & \cdot & t & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

which fixes the position of s in row k in both squares.

If $a = t$, then we can find another entry in column j_1 of A which is equal to t and is not in the first k rows of A . So, by applying Lemma 1, we can transfer the above (possibly improper) Latin square into a proper Latin square (using at most $\frac{n-1}{2}, \pm 1$ -moves) without changing the first k rows of A .

If $a \neq t$, then we can find the following patterns in B and A

$$A = \begin{array}{c|ccccc} & & j_1 & & j_3 & \\ \hline & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & \cdot & t & \cdot & r & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_1 & \cdot & u + s_t & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_2 & \cdot & \cdot & \cdot & t & \cdot \end{array} \quad B = \begin{array}{c|ccccc} & & j_1 & & j_3 & \\ \hline & \cdot & \cdot & \cdot & \cdot & \cdot \\ k & \cdot & a & \cdot & t & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_1 & \cdot & b & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ i_2 & \cdot & \cdot & \cdot & d & \cdot \end{array}$$

Since the first $k - 1$ rows of A and B are the same we must have $i_2 > k$. Therefore applying Lemma 3, interchanges t and r in row k without any other changes in row k and the first $k - 1$ rows. Applying this process (at most $n - 1$ times) produces a (proper or improper) Latin square A' whose first k rows are identical to those of B . Using Lemma 1 (and the fact that B is proper), we can transfer A' into a proper Latin square with a sequence of (at most $\frac{n-1}{2}) \pm 1$ -moves. This finishes the proof by induction. In order to transfer A to B we need to change $n - 1$ rows of A and for each row we need at most $2(n - 1)^2, \pm 1$ -moves. Therefore with at most $2(n - 1)^3, \pm 1$ -moves we can transfer A to B . ■

Remark 1 *Making moves “properly”*

In [4], they introduce moves that stay within the space of (proper) Latin squares. Such moves are called proper moves. Using Theorem 1 and a simple argument, they show that the space of (proper) Latin squares is connected under these proper moves. So we just explain what a proper move is in our notation. There are two kinds of proper moves, namely “two-rowed proper moves” and “three-rowed proper moves”. In order to define them, we first define the corresponding Latin bitrades. A two-rowed Latin bitrade is defined to be a Latin bitrade of the following form:

		j_1	j_2	\cdots	j_{r-1}	j_r	
	i_1	\cdot	\cdot	\cdots	\cdot	\cdot	
		a_b	b_c	\cdots	yz	z_a	\cdot
	i_2	\cdot	\cdot	\cdots	\cdot	\cdot	
		b_a	c_b	\cdots	zy	a_z	\cdot
		\cdot	\cdot	\cdots	\cdot	\cdot	

A three-rowed Latin bitrade is a Latin bitrade T with the following properties:

1. T has exactly three nonempty rows,
2. T is the sum of two-rowed Latin bitrades T_1 and T_2 such that there is at least one cell which is nonempty in both T_1 and T_2 .

Finally, a two-rowed proper move (resp. three-rowed proper move) means adding a two-rowed Latin bitrade (resp. three-rowed Latin bitrade) to a given Latin square provided that the result is still a Latin square.

Another set of proper moves to connect the space of all Latin squares which is similar to the ones found by Jacobson and Matthews, but certainly found independently, appears in Arthur O. Pittenger [8]. Actually Pittenger’s moves, correspond to special kinds of two-rowed and three-rowed moves, discussed above.

Remark 2 The Markov chain introduced in [4] is not known to be rapidly mixing (and thus does not have proven efficiency). Mark T. Jacobson and Peter Matthews [4] state that: “in order to use either of our Markov chains to generate almost-uniformly distributed Latin squares, we must know how rapidly the chain converges to the (uniform) stationary distribution. Of our two chains, we suspect that the “improper” one mixes more rapidly, in terms of real simulation time: executing a proper move takes time comparable to that needed to execute an equivalent sequence of ± 1 -moves; substituting an equal number of random ± 1 -moves seems likely to mix things up more.”

Remark 3 Randomly generating combinatorial objects is an important problem in combinatorics. It seems plausible to apply the ideas in this paper to attack the

same problem for some other combinatorial objects such as STS's. In fact one can define the notion of an improper STS see [1].

Acknowledgements. One of the authors (E.S.M.) thanks Amin Saberi for hospitality of a short stay in Stanford University when he introduced this problem. The other author (M.A.) was partly supported by a grant from Sharif University of Technology (Center of excellence in computational mathematics) and a grant from IPM. We also thank Masood Mortezaeefar for checking the proofs and writing down a computer program for Theorem 1.

References

- [1] Peter J. Cameron. A generalisation of t -designs. *Discrete Math.*, 309(14):4835–4842, 2009.
- [2] Nicholas J. Cavenagh. The theory and application of Latin bitrades: a survey. *Math. Slovaca*, 58(6):691–718, 2008.
- [3] Diane Donovan and E. S. Mahmoodian. An algorithm for writing any Latin interchange as a sum of intercalates. *Bull. Inst. Combin. Appl.*, 34:90–98, 2002. Corrigendum: *Bull. Inst. Combin. Appl.* 37:44, 2003.
- [4] Mark T. Jacobson and Peter Matthews. Generating uniformly distributed random Latin squares. *J. Combin. Des.*, 4(6):405–437, 1996.
- [5] A. D. Keedwell. Critical sets in Latin squares and related matters: an update. *Util. Math.*, 65:97–131, 2004.
- [6] A. A. Khanban, M. Mahdian, and E. S. Mahmoodian. A linear algebraic approach to orthogonal arrays and Latin squares. *Ars Combinatoria*, to appear.
- [7] E. S. Mahmoodian and M. S. Najafian. Possible volumes of t -($v, t + 1$) Latin trades. *Ars Combinatoria*, to appear.
- [8] Arthur O. Pittenger. Mappings of Latin squares. *Linear Algebra Appl.*, 261:251–268, 1997.
- [9] Ian M. Wanless. Cycle switches in Latin squares. *Graphs Combin.*, 20(4):545–570, 2004.